

How generation intervals shape the relationship between growth rates and reproductive numbers (世代間隔分布にもとづいた成長率と再生産数の関係)

J. Wallinga and M. Lipsitch ; (邦訳 S. Yamauchi)

2020 年 5 月 30 日 (訳)

原典 : Proceedings of Royal Society B(2007) 274, 599-604.

概要

伝染の数学モデルは、新たな感染症の制御を計画する上で非常に重要な管理ツールとなっている。このようなモデルの重要な変数は、再生産数 R である。新たに発生した感染症に対する再生産数の値は、エピソードの指数関数的成長率 r の観測値から、間接的にのみ推測することができる。再生産数を成長率に関連づけるいくつかの異なる関係式が存在するので、このような推測はあいまいであり、また、これらの関係式のいずれが新しい感染に当てはまるかは不明である。この論文では、これらの異なる関係式は、世代間隔分布の形状をどのように仮定するかということにのみ依存していることを示す。このため、世代間隔分布の形状に応じて、成長率の観測値から再生産数を推定するための適切な関係式が定まる。また、すべての世代間隔がその平均値に等しいと仮定すれば、ある成長率に対して可能な再生産数範囲の上限値が得られることを示す。さらに、実際の観測結果に等しい世代間隔分布を採用すれば、再生産数を実験的に推定することも合わせて示す。

目次

1	序論	3
2	r から R を推測する	4
2.1	ロトカ・オイラー方程式の誘導	4
2.2	再生産数 R のモーメント母関数による表示	5
3	特定の世代間隔分布	6
3.1	エピソード・モデル	6
3.2	正規分布	7
3.3	デルタ分布	7
3.4	実測による分布	8
4	再生産数を追跡する	8
5	A 形インフルエンザへの適用	9

6	検討	10
付録 A	モーメント母関数	12
A.1	数学的特性	12
A.2	生物学的意味	12
付録 B	エビデミック・モデル	13
付録 C	表	14

1 序論

過去 10 年間に於いて、健康への潜在的脅威としての感染症のエピデミックに対して、劇的に多くの注意が払われたことが見て取れる。これは、一部には家畜の伝染病発生によるものであり (Keeling ら 2001)、また天然痘ウイルスによるバイオテロ攻撃の恐れ (Gani および Leach 2001)、2003 年の重症急性呼吸器症候群 (SARS) の出現 (Lipsitch ら 2003)、そして、ヒト集団におけるインフルエンザ・パンデミックのリスク (Longini ら 2004; Ferguson ら 2005) によるものである。そのような健康に対する脅威の緩和と制御の計画は、感染伝播の数学モデルにますます依存するようになっていく。

数学的伝播モデルにおけるキーパラメータの 1 つは、再生産数 R_0 であり、完全な感受性集団内において、ある代表的な一次症例から発生する二次感染の数として定義されている。感染が一部に免疫を持つ集団内に拡散しつつあるとき、(実際の条件下において) 代表的なある一次症例から発生する二次感染の数として定義される実効再生産数 R を用いて作業する方が、多くの場合で便利である。 R の大きさは、エピデミックのリスクと感染を制御するのに要する努力との両面で有用な指標である (Anderson および May 1991; Roberts および Heesterbeek 2003; Heffernan 他 2005)。再生産数の値の正確な推定は感染の制御を計画する上で決定的に重要である。

2003 年の SARS のような新たに出現した感染に対して、新たな感染症エピデミックの伝播性に関する入手可能な情報は、新しい症例の日々の観測数に制限される場合が多い。エピデミックの初期の段階では、これらの観測値が指数関数的に増加することは、よく知られている。その指数関数的成長率 r は、単位時間における新たな感染者の、一人あたり増加率 ("per capita" change) ^{*1} として定義される。成長率 r の観測値は、線形関係式 $R = 1 + rT_c$ (Anderson および May 1991; Pybus 他 2001; Ferguson 他 2005) を通じて、再生産数 R の値に関連付けることができる。ここで T_c は平均世代間隔であり、一次感染者の感染時期と二次感染者の感染時期との間の期間長さの平均値として定義され、時には、症例間隔 (serial interval) または世代時間 (generation time) と呼ばれることもある。

人口学者、生態学者、および進化学者は、少し異なるアプローチを採用する。彼らは、いわゆるロトカ・オイラー方程式 (Lotka-Euler equation) にしたがって、年齢別出生率 (fecundity rate)、生存率 (survival rate)、および再生産数 R から成長率を導き出す (Dublin および Lotka 1925; Feller 1941; Metz および Diekmann 1986; Keyfitz および Caswell 2005)。生態学の教科書は、世代時間の変動を無視することによりこの方程式を単純化することを提案している (Begon 他 1996)。その結果は、整理すると指数関係式 $R = \exp(rT_c)$ となる。ここで、 T_c はコーホート世代時間 (cohort generation time) であり、疫学的平均世代間隔の人口学的アナロジーである。

求めるべき再生産数を成長率の観測値に関連づけるのに、2 つの相違する関係式があるので、最適のものを選択する際に困難に直面する。例えば、C 型肝炎 (Hepatitis C) のエピデミックの成長率は、年あたり $r = 0.096$ ^{*2} と見積もられていて、この感染の平均世代間隔は $T_c = 20$ 年のオーダーであるとされている (Pybus ら 2001)。線形の式による再生産数の値は $R = 1 + 0.096 \times 20 = 2.9$ であり、一方、指数式を使って得られる値は $R = \exp(0.096 \times 20) = 6.8$ である。そのような大きな不一致は、公衆衛生面での介入措置を計画する上では、大きな問題である。他に再生産数を成長率に関係付けるいくつかの表示式があり (Dublin お

^{*1} (訳注) (感染者数) $\propto e^{rt}$ と表したときの r 。

^{*2} 原文では 0.96

よび Lotka 1925; Lipsitch 他 2003; Wearing 他 2005)、また、症例数の時系列データから再生産数を推定する式 (Wallinga and Teunis 2004) がある。特定の感染について観察された成長率から再生産数を推定するため、最も適切な関係式をどのように選択すべきであろうか？

ロトカ・オイラーの方程式を人間人口学の観点から要約することから開始し、そして、Levin ら (1996) に従い、この方程式を感染症疫学のためにより便利な用語に言い換えよう。言い換えたロトカ・オイラー式を用いて、再生産数と観察された症例数変化との間のもう一つの関係の根底にある仮定を吟味する。見出した結果を、A 型インフルエンザの感染の再生産数 R を推定することにより、説明することにしよう。キーとなる変数と生態学的、人口学および疫学的用語でのその解釈は、電子補足資料 (付録) に記載される。

2 r から R を推測する

2.1 ロトカ・オイラー方程式の誘導

人口 (human population) 問題を例に、ロトカ・オイラー方程式を導く。簡単のために女性個体に焦点を当て、再生産を保証するために常に十分な男性の供給があるものと仮定する。年を単位として時間および年齢を測るものとし、現在の時間を $t = 0$ とすると、過去のイベントは t が負のときに発生し、将来のイベントは t が正のときに発生することになる。人口 (女性のみ) は一定の成長率で指数関数的に増加し、人口の年齢分布は時間的に変化しないものと仮定する。

ロトカ・オイラー方程式は、直感的に説明できる 2 つの概念の組み合わせとして理解できる。第一の概念は、ある特定の年において、あらゆる年齢の母親たちに対し、彼女らから年間生まれる子供の数を足し合わせるならば、その年の年間全出生数が得られる。 t 年において年齢 a の母親たちによる出生数は、 $t - a$ 年における出生数 (生存していない者も含めた母親の数) に、年齢 a の母親に期待される (女性) 子孫数を掛けたものに等しい。これらの出生数を母親たちに可能なすべての年齢にわたって合計すると、 t 年における全出生数が得られる。

$$b(t) = \int_{a=0}^{\infty} b(t-a)n(a)da \quad (1)$$

ここで、 $b(t)$ は t 年における全人口の出生率 (birth rate) を示し、 $n(a)$ は年齢 a の一人の母親による女性子孫の年齢別生産率 (rate of production) を示す。この方程式は、出生過程の再生方程式 (renewal equation; Feller 1941) の一例である。

第二の概念は、人口は安定した年齢分布で指数関数的に増加していると考え、任意の与えられた年 t における出生数は、 a 年前の出生数にその間の人口の指数関数的成長率を掛けたものに等しいとする。

$$b(t) = b(t-a)e^{ra} \quad (2)$$

これら 2 つの式を組み合わせると、両辺に $b(t)$ を含む次式が得られる。

$$b(t) = \int_{a=0}^{\infty} b(t)e^{-ra}n(a)da \quad (3)$$

この式 ((1) または (3)) は、次のように直感的に解釈できる。つまり、過去の各年の出生数に各個人の年数に応じた年齢別子孫生産率を掛けた値をすべて加算すると、現在の年間出生数にならねばならない。

複合パラメータ $n(a)$ は人口学分野では、生存率 (survivorship) と年齢別出生率 (fecundity function) との積、 $n(a) = l(a)m(a)$ として、より良く知られているものである。この広く知られたパラメータを用い、式の

両辺から $b(t)$ を削除して、次のロトカ・オイラー方程式が得られる。

$$1 = \int_{a=0}^{\infty} e^{-ra} n(a) da = \int_{a=0}^{\infty} e^{-ra} l(a) m(a) da \quad (4)$$

2.2 再生産数 R のモーメント母関数による表示

ロトカ・オイラー方程式は人口学の基本的部分であり、ここでは、生命表 (life table) から人口の成長率を導き出すことに興味を持たれている。一方、疫学における関連した問題は、疾病の成長率から再生産数 R を推測することである。我々は以前に、年齢 a の一人の母親による女性子孫の出生頻度として $n(a)$ を定義した。直ちにわかるように、 $n(a)$ を全寿命にわたって積分すると、一人の母親により生涯で生み出されるすべての女性子孫の数、つまり R が得られる。

$$R = \int_{a=0}^{\infty} n(a) da \quad (5)$$

一人当たり出生頻度 $n(a)$ を正規化した分布 $g(a)$ は、出産時の年齢分布となる (Metz and Diekmann 1986 を参照)。

$$g(a) = \frac{n(a)}{\int_{a=0}^{\infty} n(a) da} = \frac{n(a)}{R} \quad (6)$$

感染症の世代間隔分布 (generation interval distribution) とは、ある個体の感染からその個体による二次感染事例までの時間に対する確率分布関数である。もし感染からの経過時間を感染の「年齢」とすると、そのとき上の表記法で、世代間隔分布は上記の $g(a)$ と等価である。世代間隔分布 $g(a)$ に対するこの式をロトカ・オイラー方程式 (4) に代入すると、次式が得られる。

$$\frac{1}{R} = \int_{a=0}^{\infty} e^{-ra} g(a) da \quad (7)$$

この方程式の右辺に現れる項は、数学者には関数 $g(a)$ のいわゆるラプラス変換としてよく知られている。また特に統計学者には、それは分布 $g(a)$ のモーメント母関数 $M(z)$ として知られている (たとえば、Mood ら 1974)。

$$M(z) = \int_{a=0}^{\infty} e^{za} g(a) da \quad (8)$$

ロトカ・オイラー方程式を言い換えた式 (7) の表記を簡略化するため、このモーメント母関数を使用する。ここで、引数 z は、成長率に負号を付した値 $-r$ である：

$$R = \frac{1}{M(-r)} \quad (9)$$

ただし、 $M(-r)$ が存在するものとする。

モーメント母関数が存在するならば、それは確率分布全体の形状を一意的に特徴づけ、 $M(z)$ は $g(a)$ を決定し、逆に $g(a)$ は $M(z)$ を決定する。このモーメント母関数の表示式より得られる生物学的な結果は、成長率 r と再生産数 R の関係は世代間隔分布の形状を一意的に特徴づけるということであり、逆に、世代間隔分布の形状は再生産数と成長率との間の適切な関係を決定するということである。電子補足資料で、さらに数学的詳細を提供する。

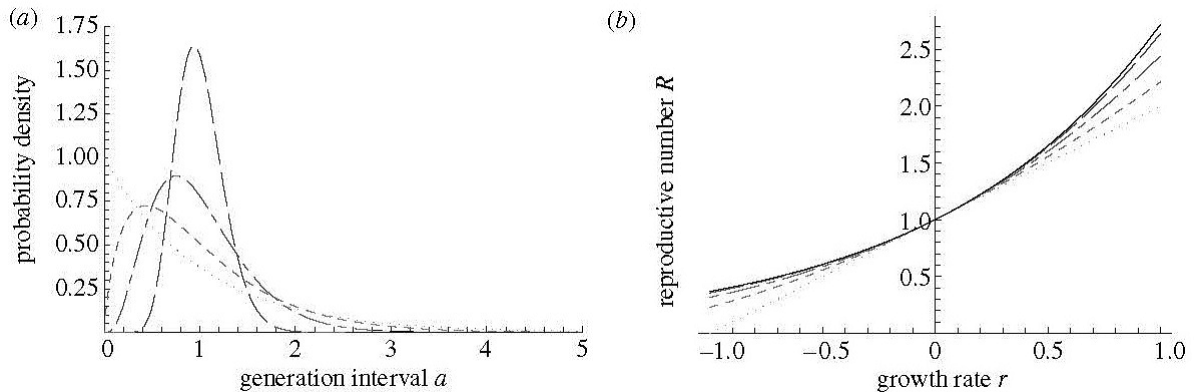


図1 世代間隔分布は、再生産数 R と成長率 r との関係を一意的に特徴づける。

(a) 世代間隔分布 $g(a)$ 。平均値 T_c を同一とし、偏差係数 σ/T_c を 0.25 (長破線)、0.5 (一点鎖線)、0.75 (短破線) から 1.0 (点線) まで増加した。

(b) 成長率 r と再生産数 R の間の対応関係。実線の曲線は、デルタ分布に対して得られる R の上限を示す。

3 特定の世代間隔分布

3.1 エピデミック・モデル

エピデミック・モデルは、暗黙裡に世代間隔分布を明示している。一つの単純なエピデミック・モデルでは、感染状況に応じて宿主の母集団を感受者 (susceptible)、感染者 (infectious)、または回復者 (recovered) のいずれかの状態に分類する。これは、いわゆるエピデミック・モデルの感受性-感染性-回復済 (SIR) クラスと呼ばれる。感染性ステージを去る時間あたり比率 (rate) は b で示され、この比率は一定であると仮定される。この感染性ステージで接触を行う頻度 (接触率) も、一定であると仮定される。世代間隔の長さは、暗黙裡に平均値 $T_c = 1/b$ の指数分布であるとして指定されている。このような指数分布は、図 1a において点線で示されている。指数分布に対するモーメント母関数 $M(z) = b/(b - z)$ に $z = -r$ を用いて (電子補足資料の表 2 を参照)、モーメント母関数式 (9) に代入する。結果を単純化すると、成長率 r と再生産数 R の間の線形関係式が得られる。

$$R = 1 + r/b \quad (10)$$

ただし、 $r > -b$ とする。成長率と再生産数の間のこの線形関係式は図 1b の点線で示されている。その線形関係は、感染症疫学で頻繁に使用され、 $1/b$ という項は平均世代間隔 (generation interval) (Ferguson ら 2005, 2006)、または感染性期間 (infectious period) の長さ (Anderson および May 1991; Pybus ら 2001) として説明されている。

エピデミック・モデルの現実性を高めるために、暴露された潜伏ステージ (感染しているがまだ感染性を持たない状態) を追加することができる。これにより、エピデミック・モデルのいわゆる感受性 (susceptible)-潜伏 (exposed)-感染性 (infectious)-回復 (recovered) のクラスが得られる (SEIR モデル)。潜伏ステージを去る単位時間あたりの割合は b_1 であり、感染ステージを去る単位時間あたりの割合は b_2 であり、両者は一定であると仮定される。これにより、世代間隔分布は 2 つの指数分布のたたみ込みとして暗黙裡に指定され、その平均値は $T_c = 1/b_1 + 1/b_2$ となる。このような分布は、右側に長い尾を引く 1 つのモードを持つ (図 1a の

短い破線を参照)。この分布のモーメント母関数 (電子補足資料を参照) を、モーメント母関数式 (9) に代入する。整理した後、次式を得る。

$$R = (1 + r/b_1)(1 + r/b_2) \quad (11)$$

ただし、 $r > \min(-b_1, -b_2)$ である。この関係式は単調増加二次曲線である (図 1b の短い破線を参照)。同じ関係式が、別のアプローチを用いて Lipsitch ら (2003) により導びかれている。

より複雑なエビデミック・モデルでは、さらに追加の潜伏段階、感染性段階を組み込んでいる。例えば、Wearing ら (2005) により提案されたエビデミック・モデルを取り上げると、それは、それぞれ b_1 の進度の x 個の潜伏段階と、それぞれ b_2 の進度の y 個の感染性段階とを有している。各段階の継続時間の分布関数から世代間隔のモーメント母関数を構成することができる (電子補足資料を参照)。得られるモーメント母関数式 (9) に代入すると、次式が得られる。

$$R = \frac{(1 + r/b_1)^x}{\sum_{i=1}^y (1 + r/b_2)^{-i}} \quad (12)$$

ただし、 $r > \min(-b_1, -b_2)$ であるとする。同じ式は Wearing ら (2005) によっても提示された。

3.2 正規分布

平均世代間隔 T_c で標準偏差 σ の感染に対して、その世代間隔はある場合には正規分布に近くなる (図 1 a の長破線)。正規分布で分布する世代間隔を仮定すると、成長率 r と再生産数 R の間に次の関係が生じる (Dublin および Lotka 1925)。

$$R = e^{rT_c - (1/2)r^2\sigma^2} \quad (13)$$

この関係は下に凸の曲線であり、近似的に指数曲線である (図 1 b の長破線)。

この式はまた、再生産数 R がより高い値になると、世代間隔は σ の値がより小さくなって、平均値の周辺により狭い範囲に集中した分布となることを示している (図 1 b を参照。より集中した分布に対応する長破線は、より分散した分布に対応する一点鎖線の上方に位置している)。

3.3 デルタ分布

すべての二次感染が正確に平均世代間隔 T_c に等しく発生する感染症に対しては、その分布はいわゆるデルタ分布となる。この分布では、モーメント母関数式 (9) は単純な指数関係に帰着する。

$$R = e^{rT_c} \quad (14)$$

この関係は指数関数的に増加する曲線である (図 1 b の実線)。

指数式 (14) は、次のように考えることにより直感的に理解することができる。つまり、時間単位 T_c 間隔での症例数の相対的な増加は、定義により $\exp(rT_c)$ であり、また、この特定のケースではすべての世代間隔は等しい長さであるため、この相対的な増加も正確に R に等しくなる。

デルタ分布は、可能な限り最も集中した分布である。成長率と平均世代間隔の所定の値に対して、再生産数は可能な限りもっとも高い値になると考えることができる。このことは以下の例で説明することができる。もし症例の半分が、平均世代間隔 T_c より少し早く二次感染を引き起こし、残りの半分がその平均より少し遅れて二次感染を引き起こすならば、より早い感染に起因する二次および三次症例の追加数は、より遅れた感染に起因する症例を補償する以上に大きくなるであろう。したがって、ある与えられた再生産数 R に対して、世

表 1 再生産数の範囲

	成長率	再生産数
減少	$r < 0$	$1 \geq R \geq e^{rT_c}$
安定	$r = 0$	$R = 1$
成長	$r > 0$	$1 < R \leq e^{rT_c}$

代間隔の持続期間に多少の変動がある伝染病は、世代間隔に変動のない伝染病よりも、より大きい成長率 r で増加するであろう。同様に、ある与えられた成長率 r に対して、世代間隔の期間に変動のない伝染病は、世代間隔にばらつきがある伝染病よりも、より高い再生産数 R で成長するであろう。

この上限の直観的な説明は、Jensen の不等式 (電子補足資料を参照) を用いた厳密な数学的論拠により支持されている。したがって、式 (14) の指数関係は再生産数が取り得る値の正確な上限を与えている。これより、成長率 r と平均世代間隔 T_c の観測値のみを使用することにより、世代間隔分布の任意の形状に対して、再生産数 R が取り得る値の範囲を示すことが可能である (表 1)。

式 (14) により再生産数を評価した場合、その過大評価の程度を相対的に表す基準を得ることもできる。式 (13) と (14) の比を取り、整理する*3。世代間隔の標準偏差 σ が 2 倍増時間 $t_d = \ln 2/r$ の 46 % 未満となる場合は、再生産数の上限と実際の値との差の相対値は常に 5 % 以下にとどまる。

3.4 実測による分布

指数関数的なエビデミックの成長期に世代間隔の期間を観察し、観察された期間のヒストグラムを用いて世代間隔分布 $g(a)$ を近似することもできる。このようなヒストグラムの区分境界を a_0, a_1, \dots, a_n で表し、これらの区分内で観測された世代間隔の相対頻度を y_1, y_2, \dots, y_n と表す。観測された分布をモーメント母関数式 (9) に代入して、積分を計算すると次式が得られる。

$$R = \frac{r}{\sum_{i=1}^n y_i (e^{-ra_{i-1}} - e^{-ra_i}) / (a_i - a_{i-1})} \quad (15)$$

観測された世代間隔のヒストグラム (図 2a) に対して、この関係式は下方に凸の増加曲線である (図 2b)。

4 再生産数を追跡する

2003 年の SARS アウトブレイクの初期フェーズのような、典型的な感染症管理状況下においては、疫学者の重要な仕事の一つは、管理措置を実施して以降の再生産数 R の変化の洞察を提供することである。このような状況下では、一定の環境の仮定や、新しい症例数の指数関数的増加などの仮定は根拠薄弱となる。

再生産数の推定値は、誕生過程の更新方程式 (式 (1)) から得ることができる。観測された全出生頻度 $\hat{b}(t)$ を $b(t)$ に用い、時間変化する感染頻度 $R_t g(a)$ を $n(a)$ に用いる。 $g(a)$ は時間 t に依存しないと仮定する。

*3 (訳注) 式で表すと

$$\frac{R \text{ の上限} - \text{実際の値}}{\text{実際の値}} = \exp\left(\frac{1}{2}r^2\sigma^2\right) - 1$$

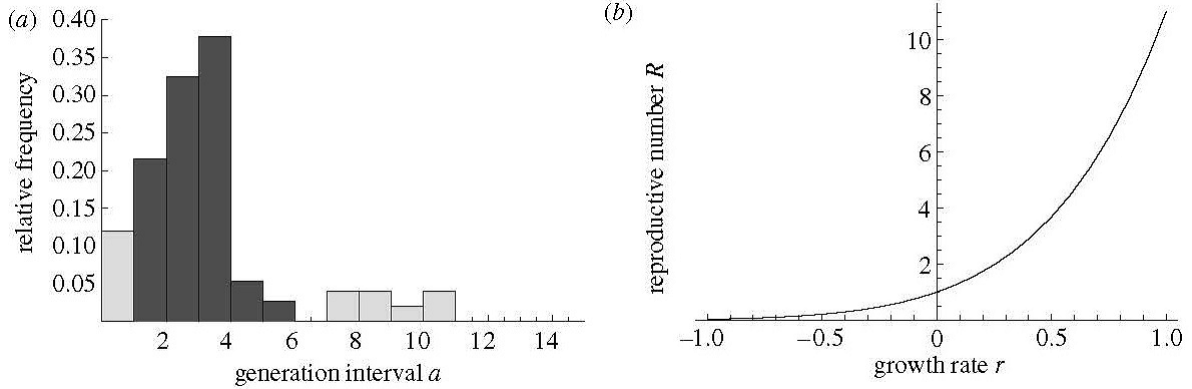


図2 A型インフルエンザ・ウイルスの感染に対して観測された世代間隔、および再生産数と成長率との間の実測的關係
 (a) 一世帯における2つの隣接したインフルエンザ症例の間で観測された世代間隔のヒストグラム (Hirtsuら2004による)。薄い網掛けは、共一次感染または三次感染による可能性のある観察を示す。
 (b) 式(15)による R と r 間の対応する実測的關係。

整理して、再生産数 R_t に対する式が得られる (C. Fraser 2006、私信より)。

$$R_t = \frac{\hat{b}(t)}{\int_{a=0}^{\infty} \hat{b}(t-a)g(a)da} \quad (16)$$

この再生産数 R_t は、二次症例が感染した時刻 t における値である。

別の方法は、一次症例が感染した時間 u における再生産数の値を用いるものである。この値は、可能な全時間 t にわたって積分することにより求められる。

$$R_u = \int_{t=u}^{\infty} g(t-u)R_t dt = \int_{t=u}^{\infty} \frac{\hat{b}(t)g(t-u)}{\int_{a=0}^{\infty} \hat{b}(t-a)g(a)da} dt \quad (17)$$

この式は、SARS の時間変化する再生産数を推定するために使用された (Wallinga および Teunis 2004)。

いずれの推定方法も、モーメント母関数の式(9)を一般化したものである。 R の表示式は、指数関数的成長曲線による世代間隔分布の積分変換に頼る一方で、 R_t と R_u の推定方法は、観察された成長曲線による世代間隔分布の積分変換に頼る。すべての場合において、再生産数の値は世代間隔分布の形状に依存する。

5 A形インフルエンザへの適用

例としてA型インフルエンザウイルスによる人の感染を取り上げ、世代間隔分布形状のさまざまな仮定が、所定の成長率に対応する再生産数の推定値へ及ぼす影響について説明する。Millsら(2004)は、アメリカの45の主要都市で、1918年のスペイン・インフルエンザの流行の初期フェーズで観察されたA形インフルエンザの成長率を分析した。彼らは、このインフルエンザ・エピデミックの初期成長率の中央値は $r = 0.20/\text{日}$ であることを見出した。

ここでは世帯内のインフルエンザ感染に焦点を当て、世帯員1人の症状発生から最初に感染した世帯員の症状発生までの期間を、世代間隔として測定する。次の2種類の観測結果を除外する必要がある。つまり、(1)最初に感染した世帯員(一次感染者)と同時に別のある世帯員が症状を発症した場合(共一次感染者)、および、(2)当該世帯員の感染が、最初に感染した世帯員による感染の可能性が低い場合(例えば第三次感染者)であ

る。日本の世帯調査で観察された A 型インフルエンザ (Hirotsu ら 2004) では、共一次感染および三次感染の可能性のある症例を除外した後、世代間隔の推定平均値 $T = 2.85$ 日、推定標準偏差 $\sigma = 0.93$ 日 となっている (図 2a を参照)。

世代間隔分布の形状について特定の仮定を行うことなく、A 型インフルエンザの再生産数は $R = 1$ より大きく、かつ $R = 1.77$ 以下であること (表 1) がわかる。推定標準偏差 $\sigma = 0.93$ 日は 2 倍増時間 $t_d = \ln 2 / 0.20 = 3.5$ 日の 46 % との基準値よりも小さいので、上限値 $R = 1.77$ はわずかながら大きく評価しているだけであることがわかる。A 型インフルエンザの再生産数の値として、SIR エピデミック・モデル (式 (10)) に対して $R = 1.57$ 、SEIR エピデミック・モデル (式 (11)) に対して $R = 1.65$ *4、1 つの潜伏ステージと 2 つの感染ステージをもつより複雑なエピデミック・モデル (式 (12)、 $x = 1, y = 2$) に対して $R = 1.66$ を得る。おそらく、最も正確な推定値は実測に基づくヒストグラム (式 (15) ; 図 2 a) を用いて得られる。これにより、A 型インフルエンザの再生産数の値は、第一次症例あたり $R = 1.73$ の第二次症例数となる。

6 検討

新しく出現した疾患では、症例数の時間的増加から観測される未知の感染因子に対して、再生産数の推論を必要とする。観察されたエピデミックの成長率と再生産数との間の既存の関係は、世代間隔分布の正確な形状に関する暗黙の仮定に関してだけ相違することを示してきた。したがって、世代間隔分布の仮定と実際の形状とが一致しているかどうかにより、観察されたエピデミック成長率と再生産数の間の関係が適切かどうかが決まる。

多くの場合、世代間隔分布の正確な形状についての知識は不十分である。我々の結果は、それにもかかわらず再生産数の値を推定することは可能であることを示している。その形状に関する情報がない場合でも、再生産数が取り得る可能な値の上限を示すことができる (表 1)。そして、世代間隔の少数の観測値を用いて、観測された成長率と再生産数の間の経験的關係式 (式 (15)) を使用できる。

再生産数は、エピデミック理論のキー概念であるとしばしば呼ばれてきた (Anderson および May 1991 ; Roberts および Heesterbeek 2003 ; Heffernan ら 2005)。再生産数の値の推定は世代間隔分布に大きく依存するにもかかわらず、この分布についてほとんど知られていないことは驚くべきことである。さまざまな基本的な疫学および生態学の問題について研究する余地がある。ここではそのうちの 3 つについて説明する。

第一に、離散的な世代時間を含めるように理論的な枠組みを拡張することができ、それは感染のモーメントが離散的な時間に結び付けられている疾病に対してより適切かも知れない。このような拡張されたフレームワークにより、広範囲の感染症および生物の再生産数の推定が可能になり、そしてそれは、生態学における生活史分析のために開発された既存の理論 (Caswell 2001) を利用することができるかもしれない。

第二に、モデルの世代間隔分布を観測した世代間隔分布に合わせることで、エピデミック・モデルのキーとなる伝播変数を推定することが可能となる。我々はインフルエンザの観測された平均世代間隔にエピデミック・モデルを適合させることで、このアプローチに従ってきた。

第三に、さまざまな感染症の世代間隔の分布を評価するために、データを収集することには利点がある。世帯内での世代間隔が異なる世帯間での世代間隔と異なるかどうかを見ることは、興味深いであろう。このような違いは、A 型インフルエンザの平均世代間隔についてのさまざまな報告の値の不一致 (Longini et al. 2004; Mills et al. 2004; Ferguson et al. 2005, 2006) を説明している可能性がある。

*4 (訳注) b_1 と b_2 に対して、 $T_c = 1/b_1 + 1/b_2$ に加えてもう一つ条件が必要である。ここでは、 $b_1 = b_2 = 2/T_c$ している。

結論として、観測した成長率を再生産数に関連づけるさまざまな式は、ロトカ・オイラーの枠組み内で理解できるであろう。その枠組みには、感染サイクルの説明と新しい症例数の変化の説明の両方が含まれる。観測された世代間隔と観測されたエビデミックの成長率とは、それらを一緒に用いると、再生産数を、したがってエビデミックを封じ込めるために必要な管理努力の適切な値を指定する。これは、必要な管理努力について医療政策立案者に知らせることを目的としている感染症サーベイランス・システムは、新たに出現する感染症の世代間隔と同様に、新しい症例の症状の発症日を監視する必要があることを意味している。

この研究は、国立総合医学研究所の感染症病原体研究モデル (協力協定 5U01GM076497 to ML)、および EC プロジェクト、INFTRANS (契約 FP6-513715) および MODELREL (契約 SANCO-790916) による支援を受けた。この論文で表明されている見解や意見は著者の意見であり、必ずしも資金提供機関のものを反映しているわけではない。この原稿の以前の版についての有用なコメントを頂いたことに関して、Tommi Asikainen、Ted Cohen、Mirjam Kretzschmar、Christina Mills、および Marcello Pagano に感謝する。

参考文献

(省略)

付録 A モーメント母関数

A.1 数学的特性

確率密度関数 $g(a)$ に従って分布する非負の確率変数 a に対して、モーメント母関数は $M_{g(a)}(z) = \int_{a=0}^{\infty} e^{za} g(a) da$ として定義される。積分が有限である場合、確率密度関数とそのモーメント母関数の間に唯一の対応関係が存在する。モーメント母関数には、多くの数学的特性がある。本文中では、下記の利用した。

1. $M_{g(a)}(z) \geq 0$ 、つまり、モーメント母関数は負値となることはない。
2. $M'_{g(a)}(z) = \int_{a=0}^{\infty} a e^{za} g(a) da > 0$ 、つまり、モーメント母関数はその引数と共に増加する。
3. $M_{g(a)}(0) = \int_{a=0}^{\infty} g(a) da = 1$ 、つまり、 $z = 0$ で評価されるモーメント母関数の値は全体の総確率を与え、定義により 1 に等しい。
4. $M'_{g(a)}(0) = \int_{a=0}^{\infty} a g(a) da = \mu$ 、つまり、 $z = 0$ でのモーメント母関数の勾配は確率変数 a の期待値を与える。
5. $\int_{a=0}^{\infty} e^{za} g(a) da \geq e^{\int_{a=0}^{\infty} a g(a) da} \Leftrightarrow M_{g(a)}(z) \geq e^{z\mu}$ 、つまり、変換された確率変数の平均は、変換が凸型である場合、これらの変数の変換された平均に少なくとも等しい (Jensen の不等式)。
6. $\ln M_{g(a)}(z) = \mu z + 1 + \frac{1}{2!} \sigma^2 z^2 + \frac{1}{3!} \sigma^3 z^3 + \dots$ 、つまり、モーメント母関数の自然対数のゼロ周りのテイラー展開は、係数として一次に確率密度関数 $g(a)$ の平均 μ 、分散 σ^2 、および歪度を持つ。
7. $g(a)$ に従って分布する確率変数 a が $f(a_1)$ および $h(a_2)$ に従って分布する 2 つの独立した確率変数 a_1 および a_2 の合計、つまり $a = a_1 + a_2$ であると考えられるとき、確率変数 a のモーメント母関数は a_1 と a_2 のモーメント母関数で構成することができる: $M_{g(a)}(z) = M_{f(a_1)}(z) \times M_{h(a_2)}(z)$ 。
8. $g(a)$ に従って分布する確率変数 a が、分布 $f(a)$ から確率 p で引き出されるものとし、分布 $h(a)$ から確率 $q = 1 - p$ で引き出されるものと考えられるとする。ただし、 g は f と h の混合である。このとき、 $g(a)$ のモーメント母関数を $f(a)$ と $h(a)$ のモーメント母関数で構成することができる: $M_{g(a)}(z) = p M_{f(a)}(z) + q M_{h(a)}(z)$ 。
9. モーメント母関数 $M_{g(a)}(z)$ は、多くのよく知られた統計分布に対して表にまとめられている。Table 2 を参照されたい。

A.2 生物学的意味

これらの数学的特性の生物学的意味を確認するために、 a で世代間隔の長さを、 $g(a)$ で世代間隔の分布を表し、その期待値 μ で平均世代間隔 T_c を表すものとする。本文で導出したように、再生産数 R は、式 $R = 1/M_{g(a)}(-r)$ にしたがって成長率 r に関連づけられる。上記のモーメント母関数の数学的特性は、再生産数 R と成長率 r の間の関係の下記の次の生物学的特性に対応する:

1. 再生産数が負になることはない。
2. 成長率が増加すると、再生産数が増加する。
3. 成長率がゼロの場合、再生産数は 1 である。
4. 成長率を再生産数に関連づける曲線の傾きは、成長率がゼロのとき、平均世代間隔 T_c に等しい。

5. 再生産数には、成長率と平均世代間隔の指数関数によって設定される上限がある： $R \leq e^{rT_c}$
6. 近似的に正規分布する世代間隔の場合、再生産数は $R \simeq e^{rT_c - \frac{1}{2}r^2\sigma^2}$ で近似される。
7. 世代間隔のモーメント母関数は、感染サイクルの引き続いて生じるステージの持続時間に対するモーメント母関数から構成することができる。たとえば、症状が発現した後でだけ感染性となる疾患を考える。引き続いて生じるステージに対して 2 つの分布を導入する。一つは $f(a)$ で、感染から感染性となるまでの期間を表し、もう一つは $h(a)$ で、感染性となってから感染させるまでの期間を与える。そのとき、複合世代間隔のモーメント母関数は $M_{g(a)} = M_{f(a)} \times M_{h(a)}$ で与えられる。
8. 世代間隔に対するモーメント母関数は、感染サイクルの (可能な複数の) 代替ループの期間に対するモーメント母関数から構成することができる。たとえば、無症候性と症候性の両症例時で感染性となるが、世代間隔の期間の分布がそれぞれの場合で異なる疾病を考える。それらの代替ループに対して 2 種類の分布を導入し、 $f(a)$ は二次感染から無症候症例の一次感染までの期間の分布を与え、 $g(a)$ は症候性症例の分布を与えるものとする。二次症例のコホートの比率 p 部分が無症候性症例に感染し、比率 q 部分が症候性症例に感染していると仮定する。複合世代間隔のモーメント母関数は、 $M_{g(a)} = pM_{f(a)} + qM_{h(a)}$ である。
9. よく知られた統計分布に従って分布する世代間隔に対して、成長率 r と再生産数 R の関係を表にまとめることができる。Table 2 を参照されたい。

付録 B エピデミック・モデル

複雑な感染サイクルに対しても、感染サイクルの引き続いて生じるイベント間の期間に対するモーメント母関数から、これらを合わせた全世代間隔に対するモーメント母関数を構成することは、比較的容易である。その構成ルールは次のとおりである。二次感染から一次感染へ戻るイベントの連鎖が 1 つしかない場合は、世代間隔に対するモーメント母関数は、その引き続いて生じるイベント間の期間に対するモーメント母関数の単なる積である。二次感染から一次感染へ戻るイベントの (可能な) 代替連鎖が複数ある場合は、世代間隔に対するモーメント母関数は、その代替連鎖のモーメント母関数の重み付き総和である。モーメント母関数の式 (9) を、感染サイクルの引き続いて生じるイベント間の期間に対するモーメント母関数の項で展開することができる。

$$R = \frac{1}{M(-r)} = \frac{1}{\sum_{i=1}^n c_i M_1(-r) \times M_2(-r) \times \cdots \times M_i(-r)}$$

ここで、 M は世代間隔全体のモーメント母関数を表し、 M_1, M_2, \dots, M_n は感染サイクルの引き続いて生じるイベント間の期間のモーメント母関数を表す。 c_i は、典型的な感染者が感染期間全体にわたって影響することにより、 i 番目のステージが二次感染の総数に及ぼす相対的な寄与度を示す。そのため、 $c_1 + c_2 + \cdots + c_n = 1$ である。この方程式を、エピデミック・モデルの再生産数を導く式のテンプレートとして使用することができる。指数分布に対するモーメント母関数 $M_i(-r) = b_i / (b_i + r)$ を 1 イベント ($n = 1, c_1 = 1$) としたこの式へ代入すると、式 (10) が得られる。同じモーメント母関数を 2 イベント ($n = 2, c_1 = 0$ および $c_2 = 1$) としたこの式に代入すると、式 (11) が得られる。モーメント母関数を $x + y$ イベント ($c_i = 0, i = 1 \sim x, c_i = 1/y, i = x + 1 \sim x + y$) のこの式へ代入すると、式 (11) が得られる。

付録 C 表

Table 1. 生態学・人口学と疫学における相似概念の表記と定義

記号	生態学および人口学	疫学	備考
$b(t)$	時間 t での単位時間出生数	時間 t での新規感染頻度	
a	年齢	感染後の時間	
$n(a)$	年齢 a の母親に予想される 女兒出生頻度 ($= l(a)m(a)$)	感染から時間 a 後の 予想される二次感染頻度	
R	再生産数 生涯で期待される子孫数	再生産数 生み出される二次感染者数	$R = \int_{a=0}^{\infty} n(a) da$ $g(a) = \frac{n(a)}{\int_{a=0}^{\infty} n(a) da} = \frac{n(a)}{R}$
$g(a)$	出生の時間分布 ($n(a)$ の正規化)	世代間隔分布 ($n(a)$ の正規化)	
r	固有の自然成長率	エピデミック成長率	
T_c	コーホート世代時間	平均世代間隔	$T_c = \int_{a=0}^{\infty} a g(a) da$
-	安定な年齢分布	感染以降の安定な年齢分布	

Table 2. 世代間隔分布、そのモーメント母関数、およびそれから生じる R と r の間の関係式

分布の種類	$g(a)$	平均	分散	$M(z)$	R	条件
指数分布	$b e^{-ba}$	$\frac{1}{b}$	$\frac{1}{b^2}$	$\frac{b}{b-z}$	$1 + \frac{r}{b}$	$r > -b$
ガンマ分布	$\frac{b^n}{\Gamma(n)} a^{n-1} e^{-ba}$	$\frac{n}{b}$	$\frac{n}{b^2}$	$\left(\frac{b}{b-z}\right)^n$	$\left(1 + \frac{r}{b}\right)^n$	$r > -b$
正規分布	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(a-\mu)^2}{2\sigma^2}}$	μ	σ^2	$e^{\mu z + \frac{1}{2}\sigma^2 z^2}$	$e^{\mu r - \frac{1}{2}\sigma^2 r^2}$	
デルタ分布	1 if $a = \mu$	μ	0	$e^{\mu z}$	$e^{\mu r}$	
均一分布	$\frac{1}{b_2-b_1}$ if $b_1 \leq a \leq b_2$	$\frac{b_1+b_2}{2}$	$\frac{(b_2-b_1)^2}{12}$	$\frac{e^{b_2 z} - e^{b_1 z}}{(b_2-b_1)z}$	$\frac{r(b_2-b_1)}{e^{-b_1 r} - e^{-b_2 r}}$	